

User-Adaptive Music Information Retrieval

Sebastian Stober, Andreas Nürnberger

Automatic structuring is one means to ease access to large music collections – be it for organisation or exploration. The AUCOMA project (Adaptive User-Centered Organization of Music Archives) aims to find ways to make such a structuring intuitively understandable to a user through automatic adaptation. This article describes the motivation of the project, discusses related work in the field of music information retrieval and presents first project results.

1 Introduction

One of the big challenges of computer science in the 21st century is the digital media explosion. Steadily growing hard-drives are filled with personal media collections comprising e.g. music, photos and videos. With increasing collection size maintenance becomes a more and more tedious task, but without manual organization effort it gets harder to access specific pieces of media or even to keep an overview. Typically, a large portion of the digital content is just “collecting dust” because the user has simply forgotten about it. Here, computer science and especially artificial intelligence can help to improve awareness and accessibility of such data: Automatic structuring is one means to ease access to media collections, be it for organization or for exploration. Moreover, users would greatly benefit if a system would not just structure the collection for easier access but would structure it in a way that is intuitively understandable for the individual user by adapting to personal preferences and needs. Unfortunately, such aspects of individualization have been only a minor issue of research in the field of Multimedia Retrieval. At best, interfaces for media collection access allow for adaptation by the user. However, they are largely lacking the ability to learn from user actions and to adapt on their own without explicit intervention of the user. The aim of the AUCOMA project is to develop intuitive, non-obstrusive, user-adaptive methods for media collection access with special focus on music information retrieval (MIR).

Dealing with music data, the following considerations serve as motivation for the project: Firstly, music can be described by a large variety of facets comprising e.g. simple tags (artist, title etc.), content-based features ranging from simple loudness to complex timbre descriptions, harmonics, meters and tempi, instrumentation, and lyrics but also information about the production and publishing process as well as the general reception in the public expressed in reviews or chart positions. This diversity of features makes music especially interesting from the data mining point of view and allows to transfer results to different domains. Secondly, perception of music is highly subjective and may depend on a person’s background. A musician, for instance, might especially look after structures, harmonics or instrumentation (possibly paying – conscious- or unconsciously – special attention to his own instrument). Non-musicians will perhaps focus more on overall timbre or general mood. Others, in turn, may have a high interest in the lyrics as long as they

are able to understand the particular language. Finally and most importantly, music can be considered an integral part of daily life even though it may often only play a background role. There may be common contexts in which music is consumed as well as contexts that are particular to an individual listener. Either way, the choice of music listened to in each context is supposed to be highly individual. The large variety of usage contexts makes MIR especially interesting for research in the area of user modelling and personalization. Given these considerations, the project’s approach to user-adaptation is two-fold:

1. Starting from various features that describe music as mentioned above, a complex multi-facet similarity measure is constructed where a similarity facet is computed either on a single feature or a combination of features. Weighting each similarity facet allows for adaptability. A weighting scheme then represents a user’s preference for grouping similar songs together and can be applied for any similarity-based structuring approach. The user is not asked explicitly to adjust the weighting to fit his needs – most likely this would be a very difficult thing to do anyway and some users might not even be aware of such preferences. Instead, the weighting has to be learned by watching the user interact with the collection. Section 2 gives details on this aspect.
2. Another way of adaptation is to utilize knowledge about personal listening habits. The idea is to infer idiosyncratic genres that represent different listening contexts. Such information can then either be used directly to browse the collection or to enrich a similarity-based structuring as orientation aid or as separate facet. Also, a combined context- and similarity-based structuring might be possible. Details on this aspect are discussed in Section 3.

2 User-Specific Similarity Measure

2.1 Adaptable and Adaptive MIR Systems

The idea of adapting similarity measures is not new: MPeer [4] allows to adjust the weight of three facets of music description in a similarity measure through an intuitive joystick interface for finding a set of similar songs given an anchor song. The facets comprise the audio content, the lyrics and cultural metadata collected from the web. From a study with 10 users, it was concluded that users tend to use nearly similar joystick settings throughout different environments.